Personal and Ubiquitous Computing, 2008

John Krumm

A Survey of Computational Location Privacy

Abstract This is a literature survey of computational location privacy, meaning computation-based privacy mechanisms that treat location data as geometric information. This definition includes privacy-preserving algorithms like anonymity and obfuscation as well as privacy-breaking algorithms that exploit the geometric nature of the data. The survey omits non-computational techniques like manually inspecting geotagged photos, and it omits techniques like encryption or access control that treat location data as general symbols. The paper reviews studies of peoples' attitudes about location privacy, computational threats on leaked location data, and computational countermeasures for mitigating these threats.

Keywords Location · Privacy · Context

1 Introduction

The temptation to give away location data grows as location-based services become more compelling. With this growth comes concern about location privacy. What are the risks if location data leaks to an unscrupulous actor? How can we avoid the bad consequences of a location leak? This paper surveys research relevant to computational location privacy, *i.e.* the ways that computation can be used to both protect and compromise location data.

Location data affords privacy techniques that exploit its geometric nature, and we limit our definition of computational location privacy to largely geometric-based algorithms. Thus we do not include protection schemes based on laws, policies, access control, standard encryption [31], and special communication protocols like mix routing [21]. Likewise, we concentrate on computational priacy attacks that take advantage of the geometric nature of location data, thus omitting attacks based on manual surveillance or hacking around standard data protection schemes.. This leaves a rich set of computational privacy attacks and countermeasures that treat location in a quantitative, geometric way. In this paper, we are most concerned with an attacker gaining access to location data and using it to algorithmically discover a subject's whereabouts and other information.

Beresford and Stajano define location privacy as

> … the ability to prevent other parties from learning one's current or past location. [6]

This definition captures the idea that the person whose location is being measured should control who can know it. It also recognizes that past location information is important to protect. While real time location could enable an attacker to find you, past data could help him or her discover who you are, where you live, and what you do. Duckham and Kulik refine the concept of location privacy by defining it as

> … a special type of information privacy which concerns the claim of individuals to determine for themselves when, how, and to what extent location information about them is communicated to others. [16]

This definition is based on Weston's definition of information privacy in general [63]. It recognizes subtle preferences of revealing location data in different forms:

- **When** – A subject may be more concerned about her current or future location being revealed than locations from the past.
- **How** – A user may be comfortable if friends can manually request his location, but may not want alerts sent automatically whenever he enters a casino or bar.
- **Extent** – A user may rather have her location reported as an ambiguous region rather than a precise point.

These different forms are the subject of many different computational schemes for protecting privacy, such as using a pseudonym instead of an actual name, intentionally adding noise to the data, and reporting location as a region instead of a point.

This paper first considers why someone would want to report their location and then reviews people's actual feelings about location privacy based on several studies. We then look at location privacy threats and countermeasures.

John Krumm
Microsoft Research, Redmond, Washington, USA
E-mail: jckrumm@microsoft.com

2 Why Reveal Location?

Location privacy would not be a problem if there were no reason to let location information leave a local device. But there are many reasons. Sometimes the very act of measuring a location involves revealing it to third parties. For example, one category of location measurement technologies use third party infrastructure to compute location, *e.g.* locations determined by cell phone providers, Google's "My Location" feature using cell tower data [20], the Loki Wi-Fi locator [42], and Quova's IP address locator [56]. These third parties perform the location computation, so they know the result. Some location-sensing systems, such as GPS, are designed to avoid this type of privacy leak by computing location on the local device. MIT's Cricket devices compute their own location indoors based on special beacons that emit radio and ultrasound [55].

Cricket and GPS are examples of so-called "inside-out" sensing in which the device to be located looks outside itself for location beacons. Such techniques have at least the potential for privacy if they can compute location on the local device as opposed to offloading the computation to a networked server. Another example is Intel's "Privacy Observant Location System" (POLS), whose name advertises one of its main features: the local device computes its own location from Wi-Fi and cell tower signal strengths using an onboard database of base station locations, eliminating the need for a third party [40] [54].

An "outside-in" location sensor depends on measurements made by the surrounding infrastructure, moving the leak risk beyond the mobile device. An example is the Ubisense indoor location system where mobile tags transmit ultra-wideband pulses to multiple, stationary receivers [62]. A central computer combines the measurements from the receivers to compute location.

Besides the act of measuring location, another reason for an individual to reveal their location is that institutions are considering the use of location data for variable pricing, pushing their users toward revealing their whereabouts. "Congestion pricing" imposes variable tolls to increase the efficient use of the road network [2]. One way to assess tolls is to examine GPS-derived location histories of drivers. "Pay as you go" insurance uses GPS measurements to determine prices based on when and where one drives.

Measured locations from vehicles are also useful as live traffic probes. This idea is used in the upcoming in-vehicle navigation systems from Dash™ and TomTom®, where users' locations will be transmitted to a server for computing real time traffic speeds.

The network architecture of location-based services sometimes requires that location measurements be transmitted to a potentially vulnerable server. For instance, a mobile user might request, from the network, bus schedules, movie times, nearby friends, discount coupons, and local information, all tied to the user's current location.

There are also several modern social applications on the Web that depend on the location of their users. Dodgeball.com has users type in their locations, which are then transmitted to a list of friends. Flickr™, and previously WWMX [60], let users geocode their photos for sharing with others. Troy's flickrvision [18] and twittervision [61] display real time, geocoded photos and micro-blog comments, respectively, on a map. Motion Based™ [48] shows and shares exercise paths recorded with GPS.

We conclude that there are applications and architectures that will entice users to reveal their location. The relative success of some location-based applications implies that at least some people are comfortable with sending their location data to third parties. The next section examines studies of peoples' willingness to share their location.

3 People Do Not Care About Location Privacy

Privacy advocates and computer scientists continue writing about location privacy, even though the general public does not show much concern about the issue. In a diary study of rendezvousing, Colbert [10] asked university students how many others they would be willing to automatically share their locations with. The students were so willing to share that the author doubted their accuracy and did not report the results. In another study of university students, Danezis *et al.* [13] asked 74 undergraduates how much they would have to be paid to share a month's worth of their location data. The median price was £10, or £20 if the data were to be used commercially. A survey of 11 participants with a mobile, location-sensitive message service found that privacy concerns were fairly light [30]. In 55 interviews with subjects in Finland, Kaasinen [32] found that "… the interviewees were not worried about privacy issues with location-aware services." However, he adds, "It did not occur to most of the interviewees that they could be located while using the service." Barkuus and Dey [5] studied 16 participants who answered questions about using imaginary, location-based services for five days. They note, "We find that people, in general, are not overly concerned about their privacy when using location-based services." In our own project to gather GPS data from drivers, we easily convinced over 250 people from our institution to give us two weeks of GPS data recorded in their car in return for a 1 in 100 chance of winning a US$ 200 MP3 player. We asked 97 of them if we could share their location data outside our institution, and only 20% said "no".

The definitive study on peoples' attitudes towards location privacy has yet to be done. Attitudes will change depending on the usefulness of the location-based service, the privacy safeguards, the amount of data, and the nature of abuses. For instance, in a study of two groups, Barkuus went on to find that the group ensconced in a closed campus environment seemed less worried about location privacy [4]. Likewise, the Danezis *et al.* study found that

the prices set on location data varied depending on who would be using it. In a wider follow-on study [12], they verified that price demands went up when the data would be used commercially. There were also increased price demands for sharing a year's worth of data compared to the original price for a month's worth.

Maybe people do not care about location privacy because they are insensitive to the negative consequences of a location leak. There *are* newspaper stories of suspicious men tracking women with GPS [51]. Another newspaper story tells how California courts have subpoenaed records from automatic toll booths on bridges. The records were used in divorce proceedings to verify claims that one spouse was not at work when he or she claimed to be [58]. In terms of sensitivity, the study in [12] found that Greeks demanded a significantly higher price for their location data than the four other EU countries surveyed. The authors note that the study followed two months after confirmed wiretapping of Greek politicians. It may be that people will become much more sensitive about location privacy if there is a major news story about the negative consequences of a location data leak. If this is true, it would be interesting to study how long this sensitivity lasts.

4 Computational Threats

The consequences of a location leak range from the uncomfortable creepiness of being watched, to unwanted revelations of a person's activities, to actual physical harm. It could be embarrassing to be seen at certain places, with the natural assumptions that follow from proximity to an abortion clinic, crack house, AIDS clinic, business competitor, or political headquarters. This has been one of the objections to Google's "Street View" geocoded images. An abortion clinic director complained that such images would compound the stress of her clients [1]. The same article reports a Street View of a man entering a pornography shop.

While the above threats come from manual inspection of leaked location data, researchers have explored what can be automatically inferred about a person based on location. This includes early work aimed at understanding peoples' movement habits, simulated privacy attacks, and sophisticated algorithms to infer context based on location. These subsections are designed to illustrate what an attacker could infer from leaked location data.

4.1 Early Analysis of Movement Patterns

Early efforts focused on extracting a person's important places from location traces, usually from GPS. Marmasse and Schmandt's commotion [44] system designated as significant those places where the GPS signal was lost three or more times within a given radius, normally due to a building blocking the signal, after which the user was prompted for a place name. Marmasse's subsequent work [43] looked at a combination of dwell time, breaks in time or distance, and periods of low GPS accuracy as potentially significant locations. Ashbrook and Starner [3] clustered places where the GPS signal was lost and asked users to name such locations. Using locations generated from Place Lab, Kang *et al.* [33] used time-based clustering to identify places that the user would likely find important. Hariharan and Toyama [23] created a time- and location-sensitive clustering technique to hierarchically represent "stays" and "destinations". Hightower *et al.*'s BeaconPrint [25] algorithm finds repeatable sets of in-range GSM and Wi-Fi base stations where a user dwells. This is interesting in that it does not use spatial coordinates as a location indicator, but instead sets of consistently heard radio transmitters. These research efforts were aimed at extracting a person's significant locations without any nefarious purpose. Other work, described next, has addressed the location privacy problem directly in an effort to assess the risk of a privacy leak.

4.2 Simulated Privacy Attacks

There are four reported studies in which researchers used recorded location data to demonstrate a privacy attack, often referred to as an "inference attack" in the literature. In three cases, the location data was pseudonomized, *i.e.* the name of the tracked person was replaced in the data with an untraceable pseudonym, as defined in [53]. Pseudonymity is a simple privacy defense against an attacker who might somehow gain access to stored location tracks. Using location measurements from their indoor Active Bat sensor, Beresford and Stajano [6] were able to find the names of all the people in their database by examining where people in the office building spent most of their time and who spent more time than anyone else at a given desk. Hoh *et al*. [27] used a database of week-long GPS traces from 239 drivers in the Detroit, MI area. Examining a subset of 65 drivers, their home-finding algorithm was able to find plausible home locations of about 85%, although the authors did not know the actual locations of the drivers' homes. A similar attack was simulated against two weeks of GPS data from 172 drivers in work by Krumm [35]. The drivers' home latitude and longitude was first determined with simple algorithms, giving a median error of about 61 meters compared to the drivers' actual home addresses. Using a reverse white pages lookup, these coordinates correctly identified 13% of the drivers' home addresses and 5% of their names.

In the fourth study demonstrating a privacy leak, Gruteser and Hoh [22] worked with GPS data that had been completely anonymized in that not even a consistent pseudonym was supplied with the time-stamped latitude and longitude coordinates. They used a standard technique from multi-target tracking [8] to accurately cluster the measured GPS points from three people. This demonstrates that even mixing together coordinates from different people is not enough to prevent an attacker from reassembling them into coherent, individual tracks. Gruteser and Hoh went on to successfully demonstrate the same attack on GPS data from five people [26].

While the Gruteser and Hoh attack above worked on GPS data, a similar approach works for anonymous indoor data. Wilson and Atkeson [64] placed simple presence sensors around a house, *i.e.* motion detectors, pressure mats, break beam sensors, and contact switches. Triggering any of these two-state sensors gave no identifying information. They developed a probabilistic tracking algorithm designed to compute which of the house's three occupants were responsible for each of the thousands of observed sensor triggers. Their algorithm made the correct association about 85% of the time, meaning they could normally tell who in the house was where.

One technique for increasing location privacy is to intentionally obfuscate the location data, possibly by adding noise or reporting regions instead of points. Duckham *et al*. [17] give a formal model of refinement operators for working around obfuscation techniques, such as assumptions about a victim's movement constraints and goal-directed behavior.

By showing how location data can be attacked with inference algorithms, researchers have established a lower bound on what an attacker could do. This research is valuable, because it highlights vulnerabilities that should be addressed.

4.3 Context Inference

Researchers have used location data to infer other things besides a person's home and work location. For example, Patterson *et al.* [52] used GPS traces to infer, in real time, a moving person's mode of transportation (*i.e.* bus, foot, car) and a prediction of their route, based on their historical behavior. Other GPS inferences include determining which road a driver is on in spite of noisy GPS data [38], and predicting where a driver will drive [19, 36] and end a trip [37]. Prediction means that privacy attacks can work on likely future locations in addition to the past and present. In some of the earliest work on the analysis of peoples' location, Newman *et al.* [49] developed PEPYS, based on an indoor location sensor in an office building. PEPYS filtered through recorded location data to find significant events like gatherings of multiple people and stopovers. 16 years later Matsuo *et al.* [46] picked up this theme and used indoor location sensing to infer many static properties about people based on where they go: age, work role, work group, work frequency, coffee drinker, smoker, work room, and which train station they used.

In light of these computational threats, Bettini *et al.* [7] declare that a sequence of recorded locations for a person constitutes a *quasi-identifier*, *i.e.* data that can be used in combination, possibly with other information, to identify the user.

Summarizing this section, location can be used to infer much about a person, even without a name attached to the data. While some of the inferences are manual, *e.g.* looking at street side images, researchers have found they can infer several personal properties automatically.

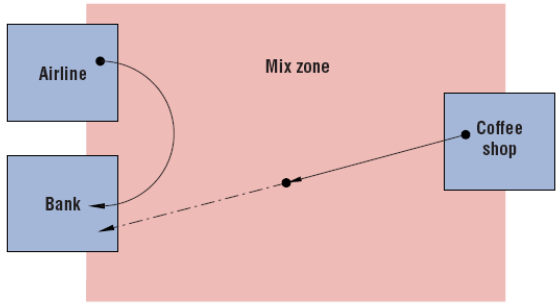


Figure 1: In a mix zone, a mobile user is given a new pseudonym to help mix him or her with other users in the same zone. Figure used by permission of IEEE, from [6].

5 Computational Countermeasures

There have been many countermeasures proposed and tested for enhancing location privacy. In their excellent survey [16], Duckham and Kulik list four general methods for ensuring location privacy:

- **Regulatory strategies** – normally government rules on the use of personal information
- **Privacy policies** – trust-based agreements between individuals and whomever is receiving their location data
- **Anonymity** – use a pseudonym and create ambiguity by grouping with other people
- **Obfuscation** – reduce the quality of the location data

In this paper we concentrate on computational countermeasures, which encompass the last two points above. Clearly a regulation or policy might specify the use of anonymity or obfuscation, but we omit a discussion of regulations and policy, concentrating instead on several different types of purely computational countermeasures.

5.1 Anonymity

Perhaps the most obvious way to maintain the privacy of location data is to replace the associated name with an untraceable ID, *i.e.* a pseudonym. A long-term pseudonym is one that persists for a given user for a long time. Beresford and Stajano [6] propose the idea of frequently changing pseudonyms for each user, which would reduce the chance that an attacker could accumulate enough history on a victim to infer their habits or identity. They note, however, that if a user's preference data were stored on a server, an attacker could link together all the pseudonyms attached to requests for a single user's data. And, as noted in Section 4, privacy researchers have demonstrated algorithms that break pseudonyms on real location data and can even reconstruct tracks from data consisting of mixed location coordinates from multiple, completely anonymized users.

Based on Sweeney's concept of *k*-anonymity for data privacy [59], Gruteser and Grunwald [21] introduced *k*-anonymity for location privacy. Instead of pseudonymously reporting his or her exact location, a person reports a region containing *k-1* other people. This presents a way of quantifying privacy: *k*-anonymity is when a person cannot be distinguished from *k-1* other people. The *k* people form the anonymity set. Gruteser and Grunwald used a quadtree to make regions, but the regions could be any shape so long as they contain *k* people, meaning an attacker could not know for sure which of the *k* people reported the location. They note that time stamps on location measurements could be similarly ambiguous.

Simple *k*-anonymity may not be enough. Bettini *et al*. [7] note that an attacker might look for patterns in the particular service requests (*e.g.* bus times, restaurant specials) that come from a pseudonymous user. These patterns could be used to link otherwise unlinkable location reports. They introduce the notion of "historical *k*-anonymity" as a way to inject ambiguity into service requests to mitigate this type of attack.

The concept of *k*-anonymity continues as a component for location privacy schemes. One particularly interesting idea, from Hashem and Kulik [24], uses one of *k-1* other mobile nodes in an *ad hoc* network to act as a query requestor in order to protect the identity of the query initiator. Mokbel *et al.* [47] developed a new query processor for location databases that lets a user specify a guaranteed *k* and the minimum area within which he or she wants to hide. Mascetti and Bettini [45] look at various ways of deciding which *k-1* reports to include, with the goal of reporting the minimum area that still contains *k* people.

The idea of changing pseudonyms for location privacy is the basis of "mix zones" of Beresford and Stajano [6]. It is assumed that people will only report their location in certain regions, called "application zones", where a location-based service is offered, *e.g.* an airport, bank, or coffee shop. In the mix zones outside the application zones, users receive new, unused pseudonyms, as shown in Figure 1. This helps prevent an attacker from linking pseudonyms, because the new pseudonym could have been assigned to anyone else in the mix zone. Beresford and Stajano show how *k* in *k*-anonymity varies with the interval between location updates for real, indoor location data taken from their institution.

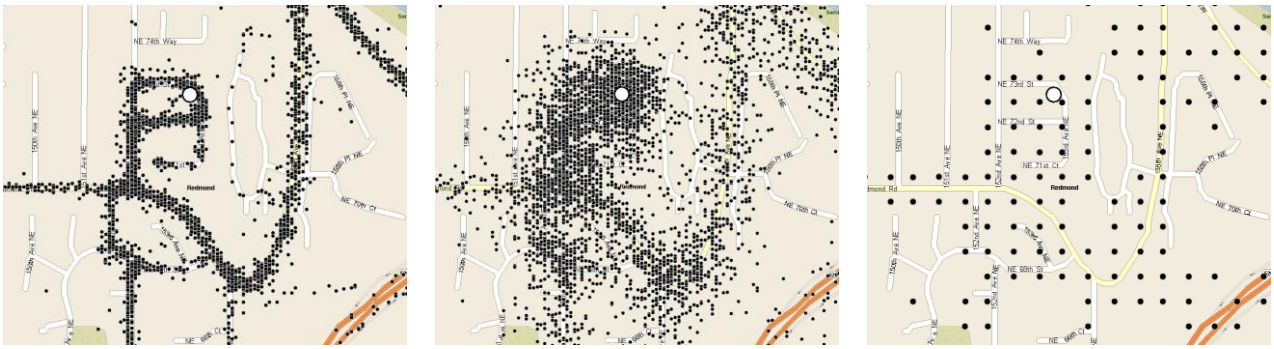
Another way to fool an attacker is to append multiple false locations to a true location report. The location-based service responds to all the reports, and the client picks out only the response based on the true location. Kido *et al*. [34] examine this technique and speculate on how to make the false locations realistic and how to reduce the cost of the inevitable extra communication. False reports are primarily useful when the user submitting them can filter out all but the server's response to a true report. Fake locations would likely render useless systems like traffic probes that use location reports to assess vehicle speeds on the road network, although recent work on the processing of "negative surveys" [29] may be a way around this problem. False reports could be tangibly expensive if they are used to trigger actions in the real world, such as turning on a light or opening a door.

5.2 Spatial and Temporal Degradation

Degrading the quality of location measurements may reduce threats to location privacy. Duckham and Kulik [14] introduce the idea of obfuscation for location privacy, formalizing the concepts of inaccuracy and imprecision as examples. Inaccuracy means giving a measurement different from the actual location, and imprecision means giving a plurality of possible locations. In an early paper on location privacy, Leonhardt and Magee discuss access control policies on location based on a hierarchy of locations, *e.g.* room, floor, building [41]. A privacy policy may specify that only higher levels of the hierarchy may be reported, leaving the user's precise location ambiguous.

Using an algorithm designed to reveal the identities of pseudonymous GPS tracks, Krumm [35] showed how much additive noise and quantization was necessary to significantly reduce the chance of a successful attack, as shown in Figure 2. The amount of obfuscation needed was surprisingly high. For instance, the inference attack still succeeded on a few subjects even after adding Gaussina noise with a standard deviation of 1 kilometer to the GPS data.

Section 4 described the multi-target tracking attack by Gruteser and Hoh [22], where a standard multi-target tracking algorithm was applied to unmix anonymous GPS tracks of multiple users. These unmixing algorithms are susceptible to mistakes where tracks appear to cross, because they cannot tell if the tracks actually crossed or just touched. Hoh and Gruteser [26] exploit this weakness in a privacy protection scheme that perturbs measured locations enough to make tracks appear to cross other peoples' tracks. Combing the idea of crossing tracks and false location reports, You *et al.* [65] demonstrate how to make false location tracks that will confuse an attacker by crossing over the true track.



(a) Original GPS data (b) Additive Gaussian noise (c) Discretized to points on grid

Figure 2: Obfuscation can enhance location privacy, *e.g.* by adding noise or discretizing to a grid. From [35].

Increasing the time between location reports can also improve privacy. In their attacks on GPS data from vehicles, Hoh *et al*. [27] show that the success rate of the attacks generally goes down with increased temporal delay between location measurements, as one would expect. Hoh *et al.* [28] go on to demonstrate an algorithm that selectively removes location samples with the goal of confusing an attacker.

5.3 Specialized Queries

Anonymity and obfuscation are general methods for location privacy that are not aimed at any certain application. There is another class of location privacy schemes intended for more specific location-based applications. Given an application, only certain types of location information might be required. A simple example is a query for a local weather report: reporting location at the resolution of a city or postal code would be sufficient, offering more privacy than GPS-level resolution.

One specialized query for location supports collaborative filtering. Users could submit ratings of where they are and receive recommendations of other places they might like. Canny presents a privacy-preserving technique for place-based collaborative filtering based on sparse factor analysis, which he shows can work directly on un-decrypted data sent in by multiple users [9].

For queries about nearby points of interest, Duckham and Kulik [15] give an algorithm for a client and server to negotiate the amount of accuracy and precision required to answer the query.

Another generic location-based service gives alerts if a friend is nearby. Louis, Lester, and Pierre [66] are three cryptography-based techniques for securely creating these alerts while guaranteeing that a friend's location will be revealed only if that friend is nearby.

Finally, one type of query asks for only aggregates of location, like the distance between two subjects, density, or average speed. While these queries seem innocent at first, a clever attacker could use results from multiple aggregate queries to infer something about an individual. Ravi *et al.* [57] have designed a query type in which an untrusted requestor submits computer code for an aggregate query (*e.g.* average). The code is first verified as "non-inference" before it can be executed on the server.

It could be the case that all location services will have a specialized location privacy technique so they can extract just what they need for the service. Perhaps sending precise coordinates, like location researchers sometimes assume, is rarely necessary. The next section briefly considers how users might configure a mix of privacy preferences.

5.4 Configuring Privacy

While it would be convenient to insert, say, a general obfuscation technique at a low level in the location stack, a gnarly "hairball" of custom-tailored techniques might emerge instead. In fact, users may demand a mix of policies. Cornwell *et al.* studied 19 participants using a people-finder application that allowed custom rules regarding location disclosure to other people, such as which groups could see the participant's location over certain time periods [11]. They found that users often set up sophisticated privacy rules and spent time revising them in response to new situations. Despite this level of care, their initial rules only satisfied the intended level of disclosure for 59% of the simulated location queries used for testing. This implies that a single location privacy mechanism may not meet a user's needs in all situations.

Configuring privacy preferences could be simplified by clustering users into coarse groups with similar privacy demands. Based on numerous opinion surveys, Westin famously grouped people into categories of privacy fundamentalists, privacy pragmatists, and privacy unconcerneds [39]. Olson *et al.* also found clusters in their study of privacy and willingness to share information [50]. They speculate that a few survey answers might be enough to determine a person's privacy category, after which their sharing preferences would be copied from other members of the same category. For location privacy, a detailed set of privacy policies for various types of location based services could be copied from a preconfigured set based on a few simple questions. These would serve as a reasonable starting point for refining a user's individual preferences for disclosing certain location data to certain people at certain times.

5.5 Quantifying Location Privacy

Progress in computational location privacy depends on the ability to quantify location privacy. There is not yet a standard for this, and it is rare for even two different research projects to use the same method of quantification. Since location can be specified as a single coordinate, one way to measure location privacy is by how much an attacker might know about this coordinate. For instance, Hoh and Gruteser [26] quantify location privacy as the expected error in distance between a person's true location and an attacker's uncertain estimates of that location.

Duckham and Kulik [15] define "level of privacy" as the number of different location coordinates sent by a user with a single location-based query. More points mean more ambiguity, and hence more privacy. The goal of their system is to be as ambiguous as possible while still getting the right answer for a point-of-interest query.

In introducing *k*-anonymity for location privacy, Gruteser and Grunwald [21] use *k* to represent the level of privacy.

Entropy is the privacy quantifier used by Beresford and Stajano [6]. They show how an attacker could use behavioral probabilities (*e.g.* a u-turn is less likely than going straight ahead) to attach probabilities to the problem of linking changing pseudonyms over time.

Hoh *et al.* [28] quantify location privacy as the duration over which an attacker could track a subject. "Time to confusion" measures how long it will take until an attacker will become confused about a subject's track as the subject seeks to obfuscate his or her location by omitting measured samples.

Of course, no matter how location privacy is quantified, it is maximized when no one knows a subject's location. This undesirable tradeoff is why many papers on computational location privacy show a relationship between location privacy and quality of service (QOS). QOS generally decreases with increasing location privacy, and researchers (*e.g.* [15, 24, 28]) seek methods of keeping both at a high level.

6 Summary

This literature survey concentrates on computational location privacy, which means computational algorithms for compromising and protecting location data. These algorithms treat location data as geometric information, not as general data. Studies show that people are generally not concerned about location privacy, although they are sensitive to how their location data could be used, and their sensitivity may rise with their awareness of privacy leaks. Researchers have demonstrated many privacy attacks on location data, including algorithms for linking pseudonyms to actual names, linking together completely anonymized location readings, and inferring much about a user based on their location history. Researchers have also proposed and tested a variety of protection schemes for leaked location data, including various types of anonymity and data obfuscation.

The research in location privacy is still young, and interesting research questions remain, such as:

- What, if anything, would trigger people to care more about location privacy? Would their concern return to its former level after the trigger?
- Can we create a complete taxonomy and formalization of computational location privacy attacks? A preliminary approach might try to show everything an attacker could infer from location data and how these inferences change with variations in the data's obfuscation level.
- If the above is possible, what is a good way to inform a user of the privacy threats of revealing location? Is there an easy-to-understand set of options for a user to choose from concerning how much they are willing to reveal?
- What is the space of applications and services that depend on location data? Given this, how much location data needs to be revealed for each of them to be successful?
- Is there a single best way to quantify location privacy? If not, is there a small set of sufficient quantifiers?

With increasing commercial interest in location based services, we expect the importance of questions like these to grow.

References

1. *Google Street Views, Cool or Creepy?*, in *New York Post*. 2007: New York, NY USA.
2. Arnott, R. and K. Small, *The Economics of Traffic Congestion*, in *American Scientist*. 1994. p. 446-455.
3. Ashbrook, D. and T. Starner, *Using GPS to Learn Significant Locations and Predict Movement across Multiple Users.* Personal and Ubiquitous Computing, 2003. **7**(5): p. 275-286.
4. Barkhuus, L., *Privacy in Location-Based Services, Concern vs. Coolness*, in *Workshop on Location System Privacy and Control, Mobile HCI 2004*. 2004: Glasgow, UK.
5. Barkuus, L. and A. Dey, *Location-Based Services for Mobile Telephony: a Study of Users' Privacy Concerns* in *9th IFIP TC13 International Conference on Human-Computer Interaction, Interact 2003*. 2003: Zurich, Switzerland. p. 709-712.
6. Beresford, A.R. and F. Stajano, *Location Privacy in Pervasive Computing*, in *IEEE Pervasive Computing Magazine*. 2003, IEEE. p. 46-55.
7. Bettini, C., X.S. Wang, and S. Jajodia, *Protecting Privacy Against Location-Based Personal Identification*, in *2nd VLDB Workshop on Secure Data Management*. 2005. p. 185-199.
8. Blackman, S.S., *Multiple-Target Tracking with Radar Applications.* 1986: Artech House.
9. Canny, J., *Some Techniques for Privacy in Ubicomp and Context-Aware Applications*, in *Workshop on Socially-informed Design of Privacy-enhancing Solutions in Ubiquitous Computing*. 2002: Goteborg, Sweden.
10. Colbert, M., *A Diary Study of Rendezvousing: Implications for Position-aware Communications for Mobile Groups*, in *ACM 2001 International Conference on Supporting Group Work*. 2001, ACM Press: Boulder, CO USA. p. 15-23.
11. Cornwell, J., et al., *User-Controllable Security and Privacy for Pervasive Computing*, in *Eighth IEEE Workshop on Mobile*

Computing Systems and Applications (HotMobile 2007). 2007: Tucson, Arizona USA.
12. Cvrček, D., et al., *A Study on The Value of Location Privacy*, in *Fifth ACM Workshop on Privacy in the Electronic Society*. 2006, ACM: Alexandria, Virginia, USA. p. 109-118.
13. Danezis, G., S. Lewis, and R. Anderson, *How Much is Location Privacy Worth?*, in *Fourth Workshop on the Economics of Information Security*. 2005: Harvard University.
14. Duckham, M. and L. Kulik, *A Formal Model of Obfuscation and Negotiation for Location Privacy*, in *3rd International Conference on Pervasive Computing (Pervasive 2005)*. 2005, Springer: Munich, Germany. p. 152-170.
15. Duckham, M. and L. Kulik, *Simulation of Obfuscation and Negotiation for Location Privacy*, in *Spatial Information Theory, International Conference, COSIT 2005*. 2005, Springer: Ellicottville, NY, USA. p. 31-48.
16. Duckham, M. and L. Kulik, *Location privacy and location-aware computing*, in *Dynamic & Mobile GIS: Investigating Change in Space and Time*, J. Drummond, et al., Editors. 2006, CRC Press: Boca Raton, FL USA. p. 34-51.
17. Duckham, M., L. Kulik, and A. Birtley, *A Spatiotemporal Model of Strategies and Counter Strategies for Location Privacy Protection*, in *4th International Conference on Geographic Information Science (GIScience 2006)*. 2006, Springer: Münster, Germany. p. 47-64.
18. flickrvision, *http://flickrvision.com/*.
19. Froehlich, J. and J. Krumm, *Route Prediction from Trip Observations*, in *Society of Automotive Engineers (SAE) 2008 World Congress*. 2008: Detroit, MI USA.
20. Google.com, *http://google.com/gmm/mylocation.html*.
21. Gruteser, M. and D. Grunwald, *Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking*, in *First ACM/USENIX International Conference on Mobile Systems, Applications, and Services (MobiSys 2003)*. 2003, ACM Press: San Francisco, CA USA. p. 31-42.
22. Gruteser, M. and B. Hoh, *On the Anonymity of Periodic Location Samples*, in *Second International Conference on Security in Pervasive Computing*. 2005: Boppard, Germany. p. 179-192.
23. Hariharan, R. and K. Toyama. *Project Lachesis: Parsing and Modeling Location Histories*. in *Third International Conference on GIScience*. 2004. Adelphi, MD.
24. Hashem, T. and L. Kulik, *Safeguarding Location Privacy in Wireless Ad-Hoc Networks*, in *9th International Conference on Ubiquitous Computing (UbiComp 2007)*. 2007: Innsbruck, Austria. p. 372-390.
25. Hightower, J., et al. *Learning and Recognizing the Places We Go*. in *UbiComp 2005: Ubiquitous Computing*. 2005.
26. Hoh, B. and M. Gruteser, *Protecting Location Privacy Through Path Confusion*, in *First International Conference on Security and Privacy for Emerging Areas in Communications Networks (SECURECOMM 2005)*. 2005, IEEE Computer Society: Athens, Greece. p. 194-205
27. Hoh, B., et al., *Enhancing Security and Privacy in Traffic-Monitoring Systems*, in *IEEE Pervasive Computing Magazine*. 2006, IEEE. p. 38-46.
28. Hoh, B., et al., *Preserving Privacy in GPS Traces via Uncertainty-Aware Path Cloaking*, in *14th ACM Conference on Computer and Communication Security (ACM CCS 2007)*. 2007: Alexandria, VA USA.
29. Horey, J., et al., *Anonymous Data Collection in Sensor Networks* in *4th Annual International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (Mobiquitous 2007)*. 2007: Philadelphia, PA USA.
30. Iachello, G., et al., *Control, Deception, and Communication: Evaluating the Deployment of a Location-Enhanced Messaging Service*, in *UbiComp 2005: Ubiquitous Computing*. 2005, Springer: Tokyo, Japan. p. 213-231.
31. Jang, Y., C. Choi, and S. Kim, *Privacy Management Mechanism for Location based Application with High Performance*, in *Communication Systems and Applications (CSA 2005)*. 2005. p. 96-101.
32. Kaasinen, E., *User needs for location-aware mobile services.* Personal and Ubiquitous Computing 2003. **7**(1): p. 70–79.
33. Kang, J.H., et al. *Extracting Places from Traces of Locations*. in *2nd ACM International Workshop on Wireless Mobile Applications and Services on WLAN Hotspots (WMASH'04)*,. 2004.
34. Kido, H., Y. Yanagisawa, and T. Satoh, *An Anonymous Communication Technique Using Dummies For Location-based Services*, in *IEEE International Conference on Pervasive Services 2005 (ICPS2005)*. 2005: Santorini, Greece. p. 88-97.
35. Krumm, J., *Inference Attacks on Location Tracks*, in *Fifth International Conference on Pervasive Computing (Pervasive 2007)*. 2007: Toronto, Ontario Canada. p. 127-143.
36. Krumm, J., *A Markov Model for Driver Turn Prediction*, in *Society of Automotive Engineers (SAE) 2008 World Congress*. 2008: Detroit, MI USA.
37. Krumm, J. and E. Horvitz, *Predestination: Inferring Destinations from Partial Trajectories*, in *UbiComp 2006: Ubiquitous Computing*. 2006: Orange County, CA USA. p. 243-260.
38. Krumm, J., J. Letchner, and E. Horvitz, *Map Matching with Travel Time Constraints (Paper 2007-01-1102)*, in *Society of Automotive Engineers (SAE) 2007 World Congress*. 2007: Detroit, MI USA.
39. Kumaraguru, P. and L.F. Cranor, *Privacy Indexes: A Survey of Westin's Studies*. 2005, School of Computer Science, Carnegie Mellon University: Pittsburgh, Pennsylvania USA. p. 22.
40. LaMarca, A., et al., *Place Lab: Device Positioning Using Radio Beacons in the Wild*, in *Third International Conference on Pervasive Computing (Pervasive 2005)*. 2005, Springer: Munich, Germany.
41. Leonhardt, U. and J. Magee, *Security Considerations for a Distributed Location Service*. Journal of Network and Systems Management, 1998. **6**(1): p. 51-70.
42. Loki, *http://www.loki.com/*.
43. Marmasse, N., *Providing Lightweight Telepresence in Mobile Communication to Enhance Collaborative Living*, in *Program in Media Arts and Sciences, School of Architecture and Planning*. 2004, MIT: Cambridge, MA, USA. p. 124.
44. Marmasse, N. and C. Schmandt. *Location-Aware Information Delivery with comMotion*. in *HUC 2K: 2nd International Symposium on Handheld and Ubiquitous Computing*. 2000. Bristol, UK: Springer.
45. Mascetti, S. and C. Bettini, *A Comparison of Spatial Generalization Algorithms for LBS Privacy Preservation*, in *International Workshop on Privacy-Aware Location-based Mobile Services (PALMS 2007)*. 2007: Mannheim, Germany.
46. Matsuo, Y., et al., *Inferring Long-term User Property based on Users' Location History*, in *20th International Joint Conference on Artificial Intelligence (IJCAI 2007)*. 2007: Hyderabad, India.
47. Mokbel, M.F., C.-Y. Chow, and W.G. Aref, *The New Casper: Query Processing for Location Services without Compromising Privacy*, in *International Conference on Very Large Data Bases (VLDB 2006)*. 2006: Seoul, South Korea. p. 763-774.
48. MotionBased, *http://www.motionbased.com/*.
49. Newman, W.M., M.A. Eldridge, and M.G. Lamming, *PEPYS: Generating Autobiographies by Automatic Tracking*, in *Second European Conference on Computer Supported Cooperative Work (ECSCW 1991)*. 1991, Springer: Amsterdam, The Netherlands. p. 175-188.
50. Olson, J.S., J. Grudin, and E. Horvitz, *A Study of Preferences for Sharing and Privacy* in *CHI '05 Extended Abstracts on Human Factors in Computing Systems*. 2005: Portland, Oregon USA. p. 1985-1988.
51. Orland, K., *Stalker Victims Should Check For GPS*, in *Associated Press*. 2003, CBS News: Milwaukee, WI USA.
52. Patterson, D.J., et al., *Inferring High-Level Behavior from Low-Level Sensors*, in *UbiComp 2003: Ubiquitous Computing*. 2003: Seattle, WA USA. p. 73-89.
53. Pfitzmann, A. and M. Köhntopp, *Anonymity, Unobservability, and Pseudonymity - A Proposal for Terminology*, in *Designing Privacy Enhancing Technologies: International Workshop on Design Issues in Anonymity and Unobservability*. 2000, Springer: Berkeley, CA USA.
54. POLS, *http://pols.sourceforge.net/*.
55. Priyantha, N.B., A. Chakraborty, and H. Balakrishnan, *The Cricket Location-Support System*, in *Sixth International Conference on Mobile Computing and Networking (MobiCom 2000)*. 2000: Boston, MA USA.
56. Quova, *http://www.quova.com/*.
57. Ravi, N., M. Gruteser, and L. Iftode, *Non-Inference: An Information Flow Control Model for Location-based Services*, in *Mobile and*

Ubiquitous Systems: Networking & Services (Mobiquitous 2006). 2006: San Jose, CA USA. p. 1-10.
58. Simerman, J., *FasTrak to Courthouse*, in *Contra Costa Times* 2007: Walnut Creek, CA.
59. Sweeney, L., *Achieving k-Anonymity Privacy Protection Using Generalization and Suppression.* International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 2002. **10**(5): p. 571-588.
60. Toyama, K., et al., *Geographic Location Tags on Digital Images*, in *11th ACM International Conference on Multimedia* 2003: Berkeley, CA USA.
61. twittervision, *http://twittervision.com/*.
62. Ubisense, *http://www.ubisense.net/*.
63. Westin, A., *Privacy and Freedom*. 1967, New York, NY USA: Atheneum. 487.
64. Wilson, D. and C. Atkeson, *Simultaneous Tracking & Activity Recognition (STAR) Using Many Anonymous, Binary Sensors*, in *Third International Conference on Pervasive Computing (Pervasive 2005)*. 2005, Springer: Munich, Germany. p. 62-79.
65. You, T.-H., W.-C. Peng, and W.-C. Lee, *Protecting Moving Trajectories with Dummies*, in *International Workshop on Privacy-Aware Location-based Mobile Services (PALMS 2007)*. 2007: Mannheim, Germany.
66. Zhong, G., I. Goldberg, and U. Hengartner, *Louis, Lester and Pierre: Three Protocols for Location Privacy*, in *7th Workshop on Privacy Enhancing Technologies*. 2007: Ottawa, Canada.